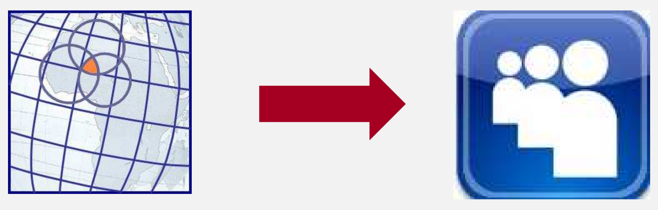
TOWARDS INTEGRATING REAL-WORLD SPATIOTEMPORAL DATA WITH SOCIAL NETWORKS

Huy Pham, Ling Hu and Cyrus Shahabi

Integrated Media Systems Center

University of Southern California

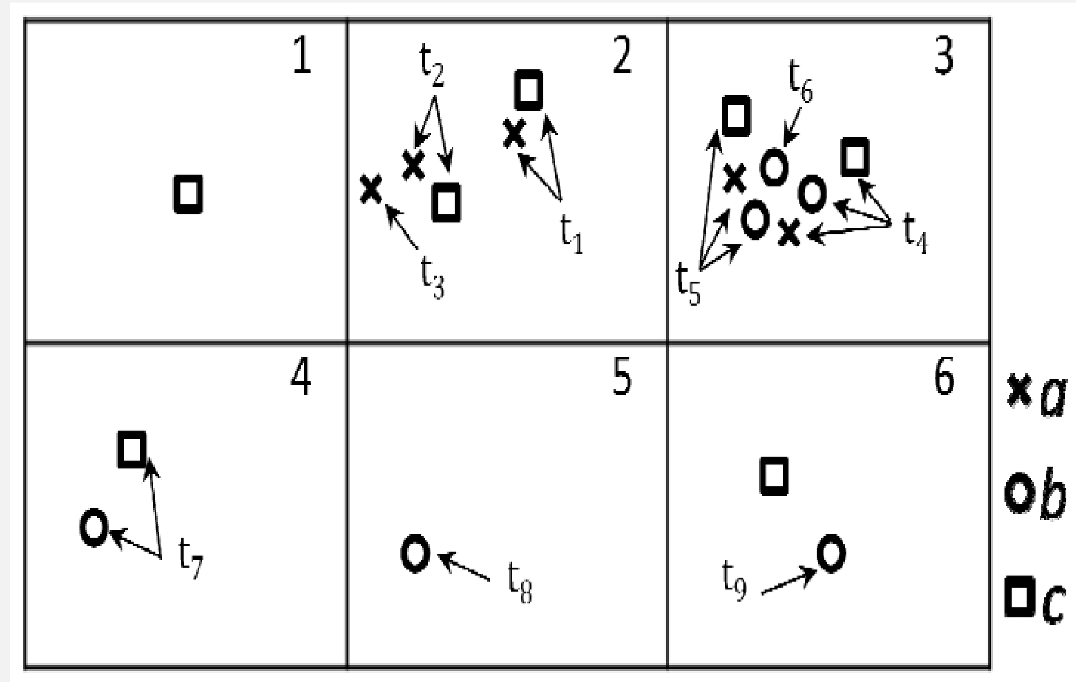
MOTIVATION



- With the popularity of GPS-embedded devices, a large amount of spatiotemporal data can be easily collected or is already available.
- Spatiotemporal data capture people's visit patterns, and the information hidden behind those data is a strong indicator of social connections.
- Necessity to seek alternative methods to infer social relationships from people's behaviors in the physical world.

PROBLEM DEFINITION

Given a set of users U, a set of places P , and triplets of social events specifying *who visited where and when*



How to infer social connections among people and how to measure the social distances quantitatively ?

GEOSO MODEL

Data Representation

Input: who visited where and when

<*user*, *location*, *time*>

Visit vectors

$V_a = (0, <t_1, t_2, t_3>, <t_4, t_5>, 0,0,0)$

$V_b = (0,0, <t_4, t_5, t_6>, t_7, t_8, t_9)$

Co-occurrence vectors

$C_{ij} = (c_{i1,j1}, c_{i2,j2}, \dots, c_{iN,jN})$

Master vector

$M = (m_1, m_2, \dots, m_k, \dots, m_N)$

$m_k = \max_{1 \le i < j \le U, 1 \le k \le N} c_{ik,jk}$

GEOSO Distance Measure

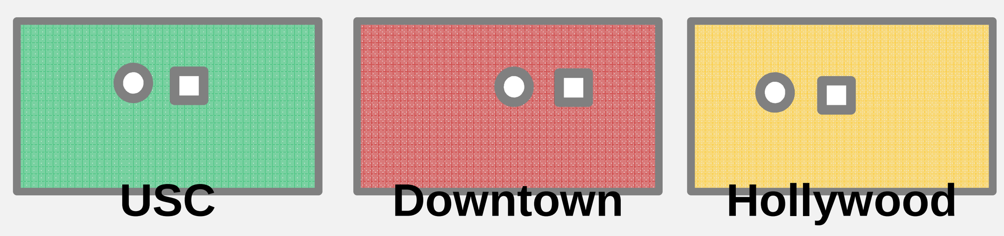
$$d_{ij} = \sqrt{\sum_k (c_{ik,jk} - m_k)^2}, \qquad s_{ij} = \frac{1}{d_{ij}}$$

Social Distance = Euclidian distance from co-occurrence vector to the master vector.

Social Similarity = Inverse of social distance.

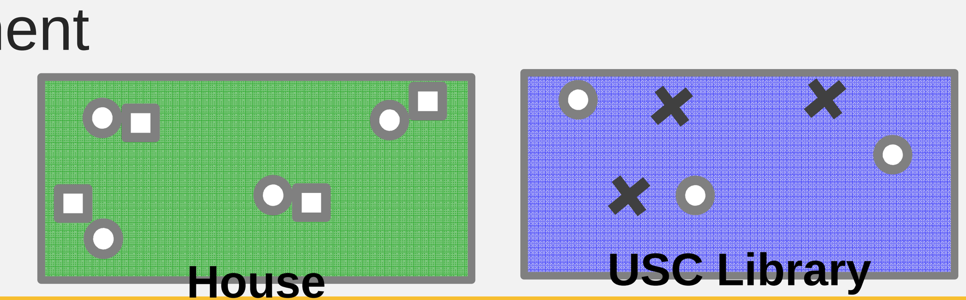
Properties

Compatibility



Two users are highly compatible

Commitment



Two users have high commitment

COMPATIBILITY V.S. COMMITMENT

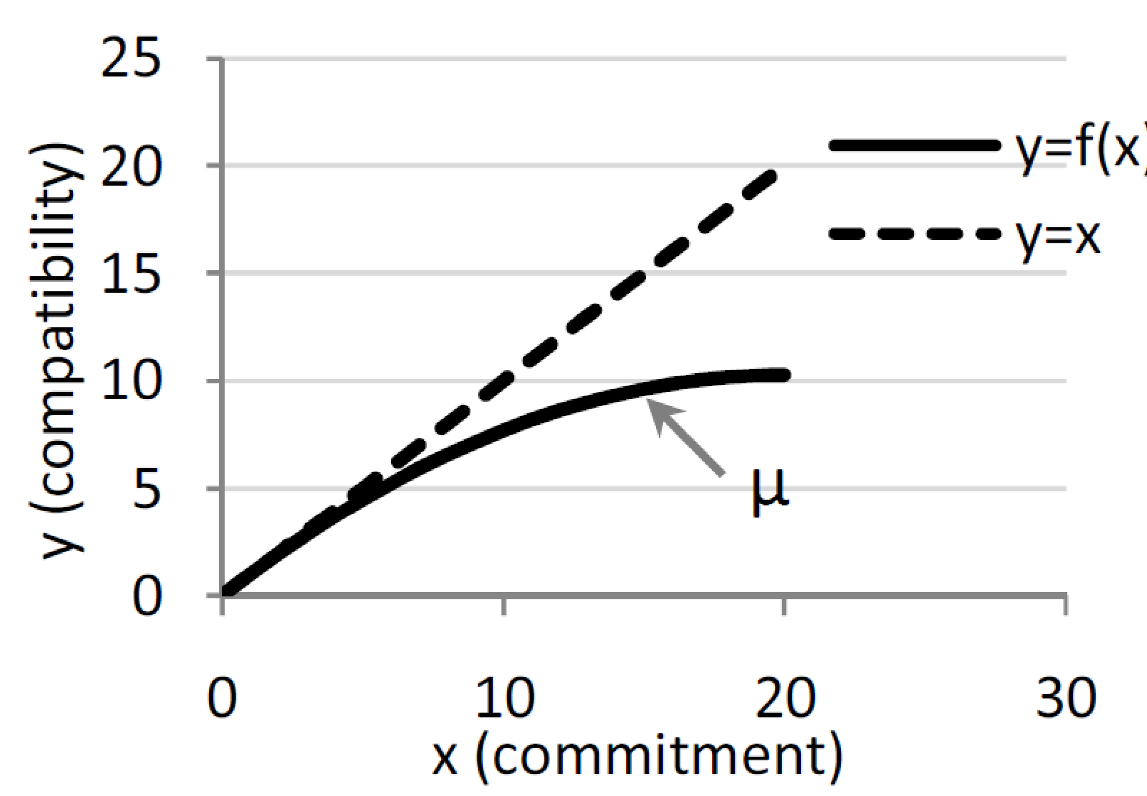
Intuitions

1. Compatibility has more impact on social distance than commitment
2. High commitment might be an indicator of coincidences

Observations

1. As the graph is located below the line y=x, compatibility calculated by GEOSO indeed has more impact on social distance than commitment.
2. As compatibility grows commitment also grows, but with a higher speed to catch up the same amount of social distance.
3. The contribution to social distance of commitment gets saturated at μ, which limits/avoids coincidences in case of high commitment.

$$y = f(x) = \frac{2mx - x^2}{2m - 1}$$



Graph of compatibility vs. commitment under the condition that they both yield the same amount of social distance.